The Annals of Statistics
2021, Vol. 49, No. 1, 255–271
https://doi.org/10.1214/20-AOS1955

WORDLENGTH ENUMERATOR FOR FRACTIONAL FACTORIAL DESIGNS

BY YU TANG[1] AND HONGQUAN XU[2]

[1]*School of Mathematical Sciences, Soochow University, ytang@suda.edu.cn*
[2]*Department of Statistics, University of California, Los Angeles, hqxu@stat.ucla.edu*

While the minimum aberration criterion is popular for selecting good designs with qualitative factors under an ANOVA model, the minimum β-aberration criterion is more suitable for selecting designs with quantitative factors under a polynomial model. In this paper, we propose the concept of wordlength enumerator to unify these two criteria. The wordlength enumerator is defined as an average similarity of contrasts among all possible pairs of runs. The wordlength enumerator is easy and fast to compute, and can be used to compare and rank designs efficiently. Based on the wordlength enumerator, we develop simple and fast methods for calculating both the generalized wordlength pattern and the β-wordlength pattern. We further obtain a lower bound of the wordlength enumerator for three-level designs and characterize the combinatorial structure of designs achieving the lower bound. Finally, we propose two methods for constructing supersaturated designs that have both generalized minimum aberration and minimum β-aberration.

1. Introduction. Fractional factorial designs are widely used in various areas for screening important factors among a large number of potential variables. The minimum aberration criterion has been frequently used in the selection of regular fractional factorial designs (Mukerjee and Wu (2006) and Wu and Hamada (2009)). It is popular especially when the experimenter has little knowledge about the potential significant effects. In order to compare general factorial designs, Tang and Deng (1999) and Xu and Wu (2001) proposed the generalized minimum aberration criterion. They further justified the criterion for designs with qualitative factors under an ANOVA model. There are abundant researches on constructions and properties of generalized minimum aberration designs; see Xu, Phoa and Wong (2009) and Cheng (2014).

For experiments with quantitative factors, response surface models such as polynomial models are frequently used for describing the relationship between the response and the factors. In such situations, the generalized minimum aberration criterion is not adequate because designs with the same generalized wordlength pattern may have very different statistical properties (Cheng and Wu (2001), Cheng and Ye (2004), Sabbaghi, Dasgupta and Wu (2014)). Cheng and Ye (2004) argued that for designs with quantitative factors, effects of lower polynomial degree should be regarded as more important than effects of higher polynomial degree whereas effects of the same polynomial degree should be regarded as equally important. Consequently, they defined the β-wordlength pattern and the minimum β-aberration criterion based on a polynomial model. Tang and Xu (2014) and Lin, Yang and Cheng (2017) provided statistical justification and additional insights regarding minimum β-aberration designs. Tang and Xu (2014) also gave some properties of the β-wordlength pattern and constructed regular minimum β-aberration designs with 27 and 81 runs. However, their method relies on the properties of regular designs, which makes it less applicable for general fractional factorial designs.

Received August 2019; revised November 2019.
MSC2020 subject classifications. 62K15.
Key words and phrases. Factor screening, generalized Hadamard matrix, generalized minimum aberration, generalized wordlength pattern, supersaturated design.

Computation is always an important issue for any algorithmic construction method because there are many nonregular designs and they do not have a unified structure (Xu, Phoa and Wong (2009)). The minimum β-aberration criterion has a major drawback in this regard. It is expensive to compute the β-wordlength pattern as it requires to consider all interaction effects by definition.

In this paper, we introduce the concept of wordlength enumerator for general fractional factorial designs. The wordlength enumerator is defined as an average similarity of contrasts among all possible pairs of runs. The wordlength enumerator is easy and fast to compute, and can be used to compare and rank designs efficiently. We show that the wordlength enumerator is a linear function of the generalized wordlength pattern for designs with qualitative factors, and a linear function of the β-wordlength pattern for designs with quantitative factors. We further establish general theoretical results which enable us to develop fast computational methods for calculating both the generalized wordlength pattern and the β-wordlength pattern. Besides the computational advantages, the wordlength enumerator offers new insights on design properties and constructions. We obtain a lower bound for wordlength enumerators and propose methods for constructing designs achieving the lower bound. The resulting designs have generalized minimum aberration as well as minimum β-aberration, whereas most existing designs do not have minimum β-aberration.

The paper is organized as follows. Section 2 introduces some notation and backgrounds. Section 3 presents the concept of the wordlength enumerator for general fractional factorial designs and studies its properties. Section 4 focuses on three-level designs and gives a lower bound of the wordlength enumerator. Section 5 proposes methods for constructing designs that achieve the lower bound. Section 6 presents conclusions and discussions. For clarity, all proofs are given in the Appendix.

2. Notation and backgrounds. A design with N runs, n factors and s levels, denoted by (N, s^n), is an $N \times n$ matrix with entries from $Z_s = \{0, 1, \dots, s-1\}$. Let $p_0(x) \equiv 1$ and $p_j(x)$ be a polynomial of degree j defined on Z_s, where $j = 1, \dots, s-1$, such that

$$\sum_{x=0}^{s-1} p_i(x) p_j(x) = \begin{cases} 0 & \text{if } i \neq j; \\ s & \text{if } i = j. \end{cases} \tag{1}$$

The set $\{p_0(x), p_1(x), \dots, p_{s-1}(x)\}$ is called an orthogonal polynomial basis (Draper and Smith (1998), Chapter 22). Denote $F_1, \dots, F_n$ as the factors of an (N, s^n) design $D = (d_{il})_{N \times n}$. The orthogonal polynomial contrast coefficient for $F_1^{j_1} \cdots F_n^{j_n}$ is defined to be an N-vector whose ith component is $p_{j_1}(d_{i1}) \times \cdots \times p_{j_n}(d_{in})$. For a number x, let $wt(x) = 0$ if $x = 0$ and $wt(x) = 1$ if $x \neq 0$. If $wt(j_1) + \cdots + wt(j_n) = k$, $F_1^{j_1} \cdots F_n^{j_n}$ is called a k-factor interaction effect as it involves k distinct factors. If $j_1 + \cdots + j_n = j$, $F_1^{j_1} \cdots F_n^{j_n}$ is also called a jth-degree interaction effect. For example, $F_1 F_2^2$ is a 2-factor interaction effect and a 3rd-degree interaction effect.

For an (N, s^n) design D, consider an ANOVA model

$$Y = X_0 \alpha_0 + X_1 \alpha_1 + \cdots + X_n \alpha_n + \epsilon,$$

where Y is the vector of N observations, α_0 is the intercept and X_0 is an $N \times 1$ vector of 1's, α_j is the vector of all j-factor interaction effects, X_j is the matrix of orthonormal contrast coefficient for α_j and ϵ is a random error. Denote $n_j = (s-1)^j \binom{n}{j}$ and $X_j = (x_{ik}^{(j)})_{N \times n_j}$, where $x_{ik}^{(j)} = p_{j_1}(d_{i1}) \times \cdots \times p_{j_n}(d_{in})$ with $wt(j_1) + \cdots + wt(j_n) = j$. Xu and Wu (2001) defined the generalized wordlength pattern of design D as

$$A_j(D) = N^{-2} \sum_{\substack{0 \le j_1, \dots, j_n \le s-1 \\ wt(j_1) + \cdots + wt(j_n) = j}} \left| \sum_{i=1}^{N} \prod_{l=1}^{n} p_{j_l}(d_{il}) \right|^2 \quad \text{for } j = 1, \dots, n. \tag{2}$$

Throughout the paper, we always define $A_0(D) = 1$. For two designs D_1 and D_2, D_1 is said to have less aberration than D_2 if there exists an $r \in \{1, 2, \ldots, n\}$, such that $A_r(D_1) < A_r(D_2)$ and $A_i(D_1) = A_i(D_2)$ for $i = 1, \ldots, r-1$. D_0 is said to have generalized minimum aberration if there is no other design with less aberration than D_0.

For an (N, s^n) design, Cheng and Ye (2004) defined the β-wordlength pattern. Consider a polynomial model

$$Y = Z_0\theta_0 + Z_1\theta_1 + \cdots + Z_K\theta_K + \epsilon,$$

where Y is the vector of N observations, θ_j is the vector of all jth-degree interactions and Z_j is the matrix of orthonormal polynomial contrast coefficient for θ_j, $K = n(s-1)$ is the highest polynomial degree, and ϵ is a random error. Denote $Z_j = (z_{ik}^{(j)})_{N\times n_j'}$ where $z_{ik}^{(j)} = p_{j_1}(d_{i1}) \times \cdots \times p_{j_n}(d_{in})$ with $j_1 + \cdots + j_n = j$ and n_j' is the number of effects with degree j. The β-wordlength pattern $(\beta_1, \ldots, \beta_K)$ is defined by

$$\beta_j(D) = N^{-2} \sum_{\substack{0\le j_1,\ldots,j_n\le s-1\\ j_1+\cdots+j_n=j}} \left| \sum_{i=1}^{N} \prod_{l=1}^{n} p_{j_l}(d_{il}) \right|^2 \quad \text{for } j = 1, \ldots, K. \tag{3}$$

Let $\beta_0(D) = 1$ for convenience. Cheng and Ye (2004) argued that a good design should sequentially minimize $\beta_1, \beta_2, \ldots, \beta_K$. Such a criterion was termed as the minimum β-aberration criterion by Tang and Xu (2014).

The $A_j(D)$ defined in (2) measures the overall aliasing of all j-factor interaction effects whereas the $\beta_j(D)$ defined in (3) measures the overall aliasing of all jth-degree interaction effects. For two-level designs, the β-wordlength pattern is the same as the generalized wordlength pattern; however, for multilevel designs with $s > 2$, the β-wordlength pattern is different from the generalized wordlength pattern.

Two designs are combinatorially isomorphic if one design can be obtained from the other by permuting rows, columns and factor levels. For geometrical isomorphism, factor level permutations are restricted to the reversal of all levels for each factor (Cheng and Ye (2004)). Geometrically isomorphic designs have the same β-wordlength pattern but combinatorially isomorphic designs may have different β-wordlength patterns.

3. Wordlength enumerator. From the expressions (2) and (3), we can easily see that both the generalized wordlength pattern and the β-wordlength pattern are defined based on relationship among columns. Although this way of thinking is statistically sound, it comes with a heavy computational burden. For an (N, s^n) design and any j, the complexity of calculating A_j or β_j is at least $O(N\binom{n}{j})$. To compute all A_j or β_j, it requires $O(Ns^n)$ operations, which is cumbersome and even unmanageable for a moderate or large n. To alleviate the computational burden, Xu and Wu (2001) used coding theory to establish a relationship between the generalized wordlength pattern and the (Hamming) distance distribution of a design. Xu (2003) further proposed an alternative criterion, called minimum moment aberration, based on the pairwise Hamming distances of a design. However, their approaches do not work for the minimum β-aberration criterion. In this section, we propose a general approach based on the concept of wordlength enumerator and develop a fast computational method for the β-wordlength pattern, which requires $O(N^2n^2(s-1))$ operations. This is crucial in the algorithmic construction of minimum β-aberration designs.

DEFINITION 1. Let $\{p_0(x), p_1(x), \ldots, p_{s-1}(x)\}$ be a set of orthogonal polynomial basis and $\{y_0, y_1, \ldots, y_{s-1}\}$ be a set of s numbers. For $u, v \in Z_s$, define the contrast similarity as

$$R(u, v) = \sum_{i=0}^{s-1} p_i(u)p_i(v)y_i. \tag{4}$$

It is easy to see that the contrast similarity $R(u, v)$ has the symmetrical property $R(u, v)=R(v, u)$ for any u and v in Z_s.

Without loss of generality, let $y_0 = 1$ throughout the paper. We can choose other y_i according to the importance of contrasts. Specifically, we consider two choices: (i) $y_i = y$ for $i = 1, \dots, s-1$ and (ii) $y_i = y^i$ for $i = 1, \dots, s-1$. The first choice assumes that all contrasts are equally important. This is appropriate for the ANOVA model with qualitative factors and the generalized minimum aberration criterion. On the other hand, the second choice assumes that lower order contrasts are more important than higher order contrasts when $0 < y < 1$. This is appropriate for quantitative factors with polynomial models and the minimum β-aberration criterion. The choice of y is arbitrary. It could be a real or complex number.

EXAMPLE 1. When $s = 3$, the orthogonal polynomials satisfying (1) are $p_0(x) = 1$, $p_1(x) = \sqrt{1.5}(x-1)$ and $p_2(x) = \sqrt{2}[1.5(x-1)^2 - 1]$ for $x \in Z_3$. It is easy to verify that

$$\begin{cases} R(0,1) = R(1,0) = R(1,2) = R(2,1) = 1 - y_2, \\ R(0,2) = R(2,0) = 1 - 1.5y_1 + 0.5y_2, \\ R(0,0) = R(2,2) = 1 + 1.5y_1 + 0.5y_2, \\ R(1,1) = 1 + 2y_2. \end{cases} \tag{5}$$

DEFINITION 2. For an (N, s^n) design $D = (d_{ik})_{N\times n}$, the wordlength enumerator of D is defined as

$$E(D) = N^{-2} \sum_{a=1}^{N} \sum_{b=1}^{N} \prod_{j=1}^{n} R(d_{aj}, d_{bj}). \tag{6}$$

The wordlength enumerator $E(D)$ is an overall measure of the contrast similarities among all possible pairs of rows in D. We use the term wordlength enumerator because $E(D)$ characterizes the generalized wordlength pattern or β-wordlength pattern when we let $y_i = y$ for $i = 1, \dots, s-1$ or $y_i = y^i$ for $i = 1, \dots, s-1$, respectively. For clarity, we denote $E(D)$ as $E_\alpha(D; y)$ or $E_\beta(D; y)$ to distinguish these two cases; see Theorems 1 and 2.

EXAMPLE 2. Consider a three-level design

$$D = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \\ 2 & 0 & 1 \end{pmatrix}.$$

For the first two rows $(0, 1, 2)$ and $(1, 2, 0)$, the product of the contrast similarities is $R(0,1)R(1,2)R(2,0) = (1 - y_2)(1 - y_2)(1 - 1.5y_1 + 0.5y_2)$. It is straightforward to verify that

$$E(D) = 1 + \left(3y_1^2 + 18y_1y_2 + 3y_2^2 + 2y_2^3 + 6y_1^2y_2\right)/4.$$

When $y_1 = y_2 = y$, we have $E_\alpha(D; y) = 1 + 6y^2 + 2y^3$. When $y_1 = y$ and $y_2 = y^2$, we have $E_\beta(D; y) = 1 + (3y^2 + 18y^3 + 9y^4 + 2y^6)/4$.

The following theorem establishes a fundamental relationship between $E_\alpha(D; y)$ and the generalized wordlength pattern under the first choice of weights.

THEOREM 1. *Let $y_i = y$ for $i = 1, \dots, s-1$. For an (N, s^n) design D,*

$$E_\alpha(D; y) = \sum_{k=0}^{n} A_k(D) y^k. \tag{7}$$

When $y_i = y$ for $i = 1, \ldots, s-1$, according to the definition of orthogonal polynomial and orthogonal matrix, we have

$$R(u, v) = \sum_{i=0}^{s-1} p_i(u) p_i(v) y_i = \begin{cases} 1 + (s-1)y & \text{if } u = v; \\ 1 - y & \text{if } u \neq v. \end{cases} \tag{8}$$

Let $d_H(a, b)$ be the Hamming distance of rows a and b, that is, the number of positions where rows a and b differ from each other. Let $B_i(D) = N^{-1}\#\{(a, b) : d_H(a, b) = i\}$ be the distance distribution for $i = 0, \ldots, n$. Then

$$\begin{aligned} E_\alpha(D; y) &= N^{-2} \sum_{a=1}^{N} \sum_{b=1}^{N} \prod_{j=1}^{n} R(d_{aj}, d_{bj}) \\ &= N^{-2} \sum_{a=1}^{N} \sum_{b=1}^{N} [1 + (s-1)y]^{n - d_H(a,b)} (1-y)^{d_H(a,b)} \\ &= N^{-1} \sum_{i=0}^{n} [1 + (s-1)y]^{n-i} (1-y)^i B_i(D). \end{aligned}$$

Combining this with Theorem 1, we get the following relationship between the generalized wordlength pattern and the distance distribution:

$$\sum_{k=0}^{n} A_k(D) y^k = N^{-1} \sum_{i=0}^{n} [1 + (s-1)y]^{n-i} (1-y)^i B_i(D). \tag{9}$$

EXAMPLE 3. Cheng and Ye (2004) considered two combinatorially isomorphic 3^{3-1} designs with different geometrical structures; see Table 1. The first design D_1 is defined by $F_3 = F_1 + F_2 \pmod 3$ and the second design D_2 is defined by $F_3' = F_1 + F_2 + 2 \pmod 3$. It is easy to verify that $B_0(D) = 1$, $B_1(D) = 0$, $B_2(D) = 6$, $B_3(D) = 2$ for both designs. Simple algebra shows that the right-hand side of (9) simplifies to $1 + 2y^3$. So we have $A_1(D) = A_2(D) = 0$ and $A_3(D) = 2$ for both designs.

Now consider the second choice of weights where $y_i = y^i$ for $i = 1, \ldots, s-1$. The situation is more complicated because (8) no longer holds. For example, we have four different $R(u, v)$ values when $s = 3$; see (5). Nevertheless, we have the following important result that relates $E_\beta(D; y)$ to the β-wordlength pattern.

TABLE 1
Two combinatorially isomorphic designs with different geometrical structures

Design D_1			Design D_2		
F_1	F_2	F_3	F_1	F_2	F_3'
0	0	0	0	0	2
0	1	1	0	1	0
0	2	2	0	2	1
1	0	1	1	0	0
1	1	2	1	1	1
1	2	0	1	2	2
2	0	2	2	0	1
2	1	0	2	1	2
2	2	1	2	2	0

THEOREM 2. *Let $y_i = y^i$ for $i = 1, \ldots, s-1$. For an (N, s^n) design D,*

$$E_\beta(D; y) = \sum_{k=0}^{n(s-1)} \beta_k(D) y^k. \tag{10}$$

EXAMPLE 4. Consider the two 3^{3-1} designs given in Table 1 again. When $y_i = y^i$ for $i = 1, 2$, it is straightforward to verify that their wordlength enumerators are

$$E_\beta(D_1; y) = 1 + (3y^3 + 3y^4 + 9y^5 + y^6)/8$$

and

$$E_\beta(D_2; y) = 1 + (3y^4 + y^6)/2,$$

respectively. According to Theorem 2, their β-wordlength patterns are $(0, 0, 3/8, 3/8, 9/8, 1/8)$ and $(0, 0, 0, 3/2, 0, 1/2)$, respectively. This agrees with the results from Cheng and Ye (2004) who computed the β-wordlength patterns according to the definition. Note that $E_\beta(D_1; y) - E_\beta(D_2; y) = 3y^3(1 - 3y + 3y^2 - y^3)/8 = 3y^3(1 - y)^3/8 > 0$ when $0 < y < 1$.

THEOREM 3. *Let D_1 and D_2 be two (N, s^n) designs. If D_1 has less β-aberration than D_2, then there exists a positive ϵ, such that for any $y \in (0, \epsilon)$,*

$$E_\beta(D_1; y) - E_\beta(D_2; y) < 0.$$

Theorem 3 implies that the wordlength enumerators can be used for identifying and ranking nonisomorphic designs. Pang and Liu (2011) and Bird and Street (2018) developed algorithms to enumerate geometrically isomorphic orthogonal arrays. Due to the complexity of isomorphism check, they only enumerated 18-run orthogonal arrays. Geometrically isomorphic designs have the same β-wordlength pattern, so they also have the same $E_\beta(D; y)$ values for all y. Whenever two designs have different $E_\beta(D; y)$ values for some y, they must be geometrically nonisomorphic. Since the computation of $E_\beta(D; y)$ is much faster than isomorphism check, one can compute and compare $E_\beta(D; y)$ to identify and rank nonisomorphic designs efficiently.

EXAMPLE 5. Pang and Liu (2011) and Bird and Street (2018) identified 13 geometrically nonisomorphic OA$(18, 3^3)$. We compute their β-wordlength patterns and $E_\beta(D; y)$ with $y = 0.005$, shown in Table 2. All 13 nonisomorphic designs have different β-wordlength patterns and different $E_\beta(D; 0.005)$ values. The ranking of $E_\beta(D; 0.005)$ values is consistent with the ranking of the β-wordlength patterns. It is certainly more efficient and convenient to compute and compare $E_\beta(D; y)$ than the whole β-wordlength pattern.

If we can obtain an explicit formula for the wordlength enumerator $E_\beta(D; y)$, then by comparing the coefficients of the monomials with the same order, we can provide closed forms for the $\beta_k(D)$ values. However, an explicit formula for the wordlength enumerator is often not available. Here, we develop an efficient method for calculating the $\beta_k(D)$ values by solving a simple linear equation system.

Let $K = n(s-1)$ and $\mathbf{B} = (\beta_1, \ldots, \beta_K)^T$ be a vector of the β-wordlength pattern. Let $\omega_1, \ldots, \omega_K$ be K distinct numbers and $E_\beta(D; \omega_j)$ be the wordlength enumerator of D evaluated at $y = \omega_j$. Denote $\mathbf{E} = (E_\beta(D; \omega_1) - 1, \ldots, E_\beta(D; \omega_K) - 1)^T$. From Theorem 2, we have $\mathbf{E} = \mathbf{V}\mathbf{B}$, where $\mathbf{V}$ is the Vandermonde matrix

$$\mathbf{V} = \begin{bmatrix} \omega_1 & \omega_1^2 & \cdots & \omega_1^K \\ \omega_2 & \omega_2^2 & \cdots & \omega_2^K \\ \vdots & \vdots & \ddots & \vdots \\ \omega_K & \omega_K^2 & \cdots & \omega_K^K \end{bmatrix}.$$

TABLE 2
Properties of geometrically nonisomorphic OA$(18, 3^3)$

Design	β_3	β_4	β_5	β_6	$E_\beta(D; 0.005) \times 10^8$
1	0	0.125	0.75	0.125	0.008
2	0	0.375	0	0.125	0.023
3	0	0.5	0	0.5	0.031
4	0	1.5	0	0.5	0.094
5	0.01	0.344	0.281	0.031	0.152
6	0.042	0.125	0.375	0.125	0.529
7	0.094	0.094	0.281	0.031	1.178
8	0.094	0.594	0.281	0.031	1.209
9	0.167	0	0	0.5	2.083
10	0.167	0.375	0	0.125	2.107
11	0.26	0.094	0.281	0.031	3.261
12	0.375	0.125	0.375	0.125	4.695
13	0.375	0.375	1.125	0.125	4.711

It is well known that the Vandermonde matrix $\mathbf{V}$ is nonsignular as long as $\omega_1, \dots, \omega_K$ are distinct, which leads to $\mathbf{B} = \mathbf{V}^{-1}\mathbf{E}$. We can choose arbitrarily K distinct numbers for $\omega_1, \dots, \omega_K$. For easy computation, we can choose $\omega_1, \dots, \omega_K$ such that $\mathbf{V}^{-1}$ has a simple closed form. Specifically, let $\omega = e^{2\pi\sqrt{-1}/K}$ and $\omega_i = \omega^i$ for $i = 1, \dots, K$ so that $\omega^K = 1$ and $\omega_i^K = 1$. Then $\mathbf{V}\overline{\mathbf{V}}^T = \overline{\mathbf{V}}^T\mathbf{V} = K\mathbf{I}$ and $\mathbf{V}^{-1} = K^{-1}\overline{\mathbf{V}}^T$, where $\overline{\mathbf{V}}^T$ is the conjugate transpose of $\mathbf{V}$ and $\mathbf{I}$ is a $K \times K$ identity matrix. Note that the (i, j)th elements of $\mathbf{V}$ and $\overline{\mathbf{V}}^T$ are ω^{ij} and ω^{-ij}, respectively. We have the following theorem.

THEOREM 4. *Let $K = n(s-1)$ and $\omega = e^{2\pi\sqrt{-1}/K}$. For an (N, s^n) design D,*

$$\beta_i(D) = \frac{1}{K}\sum_{j=1}^{K} \omega^{-ij}\big(E_\beta(D; \omega^j) - 1\big) \quad \text{for } i = 1, \dots, K. \tag{11}$$

Theorem 4 provides a simple and fast method for computing the β-wordlength pattern, which is extremely useful in practice. For an (N, s^n) design, the complexity of computing each wordlength enumerator $E_\beta(D; \omega^j)$ according to (6) is $O(N^2 n)$. So the complexity of computing the entire β-wordlength pattern according to Theorem 4 is $O(N^2 nK + K^2) = O(N^2 n^2(s-1) + n^2(s-1)^2)$, which is equivalent to $O(N^2 n^2(s-1))$. In contrast, the complexity of computing the entire β-wordlength pattern according to the definition (3) is $O(Ns^n)$. The difference is huge for moderate to large n.

EXAMPLE 6. Consider an OA$(36, 3^{13})$ listed on Sloane (2019) and given in Appendix B. Using Theorem 4, we obtain the β-wordlength pattern as $(\beta_1, \dots, \beta_{26}) = (0, 0, 7.875, 53.039, 137.426, \dots, 1.545)$. It took 4.5 seconds to compute the entire β-wordlength pattern using Theorem 4 using R on an iMac computer with 3.2 GHz Intel Core i5 processor. In comparison, it took 28, 67 and 204 seconds to compute the first three, four and five β_i according to the definition (3).

To further illustrate the use of Theorems 3 and 4, we construct minimum β-aberration designs from the OA$(36, 3^{13})$ given in Appendix B. Owing to the high efficiency of the method for calculating $\beta_k(D)$ values stated above, we adopt an exhaustive search by considering

TABLE 3
Minimum β-aberration designs from an OA(36, 3^{13}) *in Appendix* B

n	Columns and level permutations									(β_3, β_4)
3	$3u$	4	13							(0, 0.0313)
4	2	3	$8u$	$13u$						(0, 0.4219)
5	1	$2u^2$	6	7	13					(0, 2.875)
6	1	$2u^2$	$3u$	6	7	13				(0.0938, 4.1016)
7	1	$4u^2$	5	$8u$	$10u^2$	11	12			(0.2813, 5.8125)
8	$1u$	2	$5u^2$	$6u$	$7u$	8	12	13		(0.6563, 11.9453)
9	$1u$	2	$3u$	$4u^2$	$5u^2$	$7u$	$9u$	12	13	(1.1719, 17.7891)
10	$1u^2$	$2u$	3	$4u^2$	5	$7u^2$	$8u^2$	$10u^2$	11	(1.875, 21.375)
	12									
11	$1u^2$	$2u$	3	$4u^2$	5	6	$7u^2$	$8u^2$	$10u^2$	(2.8125, 29.8828)
	11	12								
12	$1u^2$	$2u$	3	$4u^2$	5	6	$7u$	$8u^2$	9	(3.75, 44.5313)
	$10u^2$	11	12							
13	$1u^2$	$2u$	3	$4u^2$	5	6	$7u$	$8u^2$	9	(5.6719, 61.7578)
	$10u^2$	11	12	13						

Note: u and u^2 represent $\{0, 1, 2\} \to \{1, 2, 0\}$ and $\{0, 1, 2\} \to \{2, 0, 1\}$, respectively.

all projections and conducting all level permutations to find the minimum β-aberration designs. There are six level permutations for three levels; however, as pointed by Cheng and Ye (2004), the six permutations can be divided into three pairs when geometrical isomorphism is considered. Following Cheng and Ye (2004), we only need to consider three linear permutations, that is, the identical transformation I: $\{0, 1, 2\} \to \{0, 1, 2\}$, u: $\{0, 1, 2\} \to \{1, 2, 0\}$ and u^2: $\{0, 1, 2\} \to \{2, 0, 1\}$. For example, when we consider 10-factor projections, there are $\binom{13}{10} = 286$ projections. For each projection, there are $3^{10} = 59{,}049$ linear level permutations. For each of these designs, we compute the β-wordlength pattern according to Theorem 4. If the original formula of β-wordlength pattern (3) is used, the computational time needed is intolerable. Table 3 lists the columns and level permutations of the best projections with $n = 3$–13 columns. According to Table 3, for $n = 10$, to obtain the minimum β-aberration projection design, we should choose columns 1–5, 7–8 and 10–12, and further apply level permutation u to column 2, level permutation u^2 to columns 1, 4, 7, 8 and 10. The resulting design has $\beta_3 = 1.875$ and $\beta_4 = 21.375$.

We can also develop a fast computational procedure for calculating the generalized wordlength pattern $(A_1, \dots, A_n)$ based on the wordlength enumerator $E_\alpha(D; y)$ with a complexity of $O(N^2n^2)$ for (N, s^n) designs. The procedure is similar to Theorem 4 so we omit the details.

4. A lower bound for three-level designs. In the rest of the paper, we focus on three-level designs, which are commonly used for studying quantitative factors. For an $(N, 3^n)$ design $D = (d_{il})_{N\times n}$ and $a, b = 1, \dots, N$, let

$$
\begin{cases}
n_1(a, b) = \sharp\{l : (d_{al}, d_{bl}) = (0, 1), (1, 0), (1, 2) \text{ or } (2, 1), l = 1, \dots, n\},\\
n_2(a, b) = \sharp\{l : (d_{al}, d_{bl}) = (2, 0) \text{ or } (0, 2), l = 1, \dots, n\},\\
n_3(a, b) = \sharp\{l : (d_{al}, d_{bl}) = (0, 0) \text{ or } (2, 2), l = 1, \dots, n\},\\
n_4(a, b) = \sharp\{l : (d_{al}, d_{bl}) = (1, 1), l = 1, \dots, n\}.
\end{cases}
\tag{12}
$$

For convenience, denote

$$
\sigma_1 = 1 - y_2, \qquad \sigma_2 = 1 - 1.5y_1 + 0.5y_2, \qquad \sigma_3 = 1 + 1.5y_1 + 0.5y_2, \qquad \sigma_4 = 1 + 2y_2.
$$

Following (5) and (6), we have

$$E(D) = N^{-2} \sum_{a=1}^{N} \sum_{b=1}^{N} \sigma_1^{n_1(a,b)} \sigma_2^{n_2(a,b)} \sigma_3^{n_3(a,b)} \sigma_4^{n_4(a,b)}. \tag{13}$$

THEOREM 5. *Suppose $\sigma_i > 0$ for $i = 1, \dots, 4$. For a balanced $(N, 3^n)$ design D,*

$$E(D) \geq N^{-1}\left[\left(\sigma_3^2 \sigma_4\right)^{n/3} + (N-1)\left(\sigma_1^2 \sigma_2\right)^{\delta}\left(\sigma_3^2 \sigma_4\right)^{n/3-\delta}\right], \tag{14}$$

where $\delta = 2nN/[9(N-1)]$. The above lower bound can be achieved if and only if the following conditions are satisfied (i) *$n_4(a,a) = n/3$ for all rows a and* (ii) *$n_1(a,b) = 2\delta$, $n_2(a,b) = \delta$, $n_3(a,b) = 2(n/3 - \delta)$ and $n_4(a,b) = n/3 - \delta$ for all distinct rows a and b, where $n_1(a,b)$, $n_2(a,b)$, $n_3(a,b)$ and $n_4(a,b)$ are defined in* (12).

Substituting $y_1 = y_2 = y$ or $y_1 = y$, $y_2 = y^2$ into (14), respectively, and collecting like terms, we obtain the following corollary.

COROLLARY 1. *For a balanced $(N, 3^n)$ design D and $0 < y < 1$,*

$$E_\alpha(D; y) \geq 1 + \frac{n(2n - N + 1)}{N - 1} y^2 + \cdots + \frac{2^n}{N}\left[1 + (N-1)(-2)^{\frac{-2nN}{3(N-1)}}\right] y^n,$$

and

$$E_\beta(D; y) \geq 1 + \frac{n(n - N + 1)}{2(N - 1)} y^2 + \cdots + 2^{-n/3} y^{2n}.$$

For a balanced $(N, 3^n)$ design D, both $A_1(D)$ and $\beta_1(D)$ are zero. Then Corollary 1 provides two lower bounds on $A_2(D)$ and $\beta_2(D)$ as y goes to 0. Specifically, we have

$$A_2(D) \geq \frac{n(2n - N + 1)}{N - 1} \quad \text{and} \quad \beta_2(D) \geq \frac{n(n - N + 1)}{2(N - 1)}. \tag{15}$$

A balanced three-level design D achieving the lower bound in Theorem 5 has generalized minimum aberration and minimum β-aberration among all possible designs. Theorem 5 characterizes the combinatorial structure of such designs. The condition $n_4(a,a) = n/3$ is equivalent to that each row contains $n/3$ ones. The conditions on $n_i(a,b)$ ensure D to have proper level balances for all possible pairs of rows. In addition, such a design is equidistant. It is easy to see that the L_1- and L_2-distances between distinct rows of a and b are 4δ and $\sqrt{6\delta}$, respectively. Then, by Theorem 3 of Zhou and Xu (2015), D is a maximin distance design under both the L_1 and L_2 distances.

EXAMPLE 7. The wordlength enumerators of the $(9, 3^{12})$ design in Table 4 are

$$E_\alpha(D; y) = 1 + 24y^2 + 224y^3 + \cdots + 448y^{12},$$
$$E_\beta(D; y) = 1 + 3y^2 + 45y^3 + \cdots + 0.0625y^{24},$$

respectively. So we have $A_2(D) = 24$ and $\beta_2(D) = 3$, achieving the lower bounds of $A_2(D)$ and $\beta_2(D)$ in (15), respectively. It is easy to verify that the conditions for achieving the lower bound (14) in Theorem 5 are satisfied, so the design has generalized minimum aberration and minimum β-aberration.

TABLE 4
An optimal $(9, 3^{12})$ *design*

Run	1	2	3	4	5	6	7	8	9	10	11	12
1	0	0	0	0	1	1	1	1	2	2	2	2
2	0	1	1	2	1	2	2	0	2	0	0	1
3	0	2	2	1	1	0	0	2	2	1	1	0
4	1	0	1	1	2	1	2	2	0	2	0	0
5	1	1	2	0	2	2	0	1	0	0	1	2
6	1	2	0	2	2	0	1	0	0	1	2	1
7	2	0	2	2	0	1	0	0	1	2	1	1
8	2	1	0	1	0	2	1	2	1	0	2	0
9	2	2	1	0	0	0	2	1	1	1	0	2

5. Construction of optimal supersaturated designs. We present two construction methods for three-level designs achieving the lower bound of $E(D)$ in Theorem 5. The resulting designs are supersaturated in the sense that the number of factors is larger than the number of runs. Supersaturated designs are useful for screening a large number of factors using a small number of runs. There are many existing researches on the construction and analysis of such designs after Lin (1993) and Wu (1993); see, for example, Sun, Lin and Liu (2011) and Georgiou (2014) for a review.

Our first construction method uses saturated regular designs. We have the following general result.

THEOREM 6. *For any integer $k \geq 2$, there exists an optimal $(N, 3^n)$ design with $N = 3^k$ and $n = 3(3^k - 1)/2$ achieving the lower bound in Theorem 5.*

EXAMPLE 8. A regular 3^{4-2} design has two independent columns F_1 and F_2 and two dependent columns F_3 and F_4, where $F_3 = F_1 + F_2 \pmod 3$ and $F_4 = F_1 + 2F_2 \pmod 3$. Then the following 12 columns,

$$(F_1, F_2, F_3, F_4, F_1 + 1, F_2 + 1, F_3 + 1, F_4 + 1, F_1 + 2, F_2 + 2, F_3 + 2, F_4 + 2),$$

form an optimal $(9, 3^{12})$ design, which is listed in Table 4, where the addition is over Z_3.

Xu and Wu (2005) presented several methods for constructing optimal multilevel supersaturated designs. Their designs have generalized minimum aberration, but do not have minimum β-aberration. The designs we constructed from Theorem 6 have both generalized minimum aberration and minimum β-aberration. Here is an example.

EXAMPLE 9. Xu and Wu (2005) provided two constructions of $(9, 3^{12})$ designs via their Theorems 6 and 7. Denote these two designs by D_3 and D_4 for convenience. Both designs have the same $E_\alpha(D; y)$ as the optimal $(9, 3^{12})$ design given in Table 4; see Example 7. Their $E_\beta(D; y)$ are different and they are

$$E_\beta(D_3; y) = 1 + 6.5y^2 + 41.625y^3 + \cdots + 0.1292y^{24},$$
$$E_\beta(D_4; y) = 1 + 7.5y^2 + 31.5y^3 + \cdots + 0.1526y^{24},$$

respectively. For $y = 0.001$, $E_\beta(D_3; y) = 6.5418 \times 10^{-6}$ and $E_\beta(D_4; y) = 7.5317 \times 10^{-6}$. The lower bound of $E_\beta(D; y)$ given in Theorem 5 is 3.0451×10^{-6}, which is achieved by the optimal $(9, 3^{12})$ design given in Table 4.

In addition, our new design is better than the designs from Xu and Wu (2005) in terms of other aspects such as the number of nonorthogonal pairs and maximum absolute correlation between the columns. Our optimal design has 12 nonorthogonal pairs, each with absolute correlation 0.5, whereas the two designs from Xu and Wu (2005) have 38 and 54 nonorthogonal pairs, respectively, and maximum absolute correlation 0.67 for both.

Huang, Lin and Liu (2012) proposed the γ-wordlength pattern to characterize supersaturated designs with quantitative factors. Such a criterion only considers the main effects and two-factor interactions, thus can be regarded as a simpler version of the β-wordlength pattern. In Example 4 of their paper, the authors listed a 9×28 array, which consists of seven separated $\mathrm{OA}(9, 3^4)$s. We consider $(9, 3^{12})$ designs formed by three $\mathrm{OA}(9, 3^4)$s. There are $\binom{7}{3} = 35$ such $(9, 3^{12})$ designs. For each design, we conduct all level permutations to find the best designs. The 35 designs fall into two groups. One group has 21 designs which can achieve the β-wordlength pattern of $(0, 4, 44.25, \dots, 0.0039)$ or the γ-wordlength pattern of $(0, 4, 16, 4)$. Using the notation in Huang, Lin and Liu (2012), one example of such designs is $1\,2\,3$ $(acac\,baaa\,cbac)$. The other group has 14 designs which can only achieve the β-wordlength pattern of $(0, 4.25, 42.375, \dots, 0.015625)$ or the γ-wordlength pattern of $(0, 4.25, 15.5, 4.25)$. One such design is $1\,2\,4$ $(aaca\,cacb\,bbcb)$. When $y = 0.001$, $E_\beta(D; y) = 4.0444 \times 10^{-6}$ or 4.2925×10^{-6}, respectively, for these two designs. Both designs do not achieve the lower bound of $E_\beta(D; y)$ given in Theorem 5 or Corollary 1.

Our second construction method uses generalized Hadamard matrices. A generalized Hadamard matrix over an additive group G of order s, denoted by $H(\lambda, G) = (h_{ij})$, is a $(\lambda s) \times (\lambda s)$ matrix with entries from G satisfying that for every i, j, $1 \le i < j \le \lambda s$, the multiset $\{h_{ik} - h_{jk} | 1 \le k \le \lambda s\}$ contains every element of G exactly λ times. A generalized Hadamard matrix is normalized if all entries in the first row and first column of the matrix are zero. When $s = 2$, a generalized Hadamard matrix is nothing but a Hadamard matrix.

EXAMPLE 10. The following 6×6 matrix is a generalized Hadamard matrix $H(2, Z_3)$:

$$\begin{matrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 2 & 2 \\ 0 & 1 & 0 & 2 & 1 & 2 \\ 0 & 1 & 2 & 0 & 2 & 1 \\ 0 & 2 & 1 & 2 & 0 & 1 \\ 0 & 2 & 2 & 1 & 1 & 0 \end{matrix}$$

Let $H_0 = (h_{ij})$ be the 6×5 matrix obtained by deleting the first column. Let $H_1 = ((h_{ij} + 1) \mod 3)$ and $H_2 = ((h_{ij} + 2) \mod 3)$. Then the column juxtaposition of H_0, H_1 and H_2 forms the $(6, 3^{15})$ design D given in Table 5, which achieves the lower bound in Theorem 5.

TABLE 5
An optimal $(6, 3^{15})$ design

Run	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0	0	0	0	0	1	1	1	1	1	2	2	2	2	2
2	0	1	1	2	2	1	2	2	0	0	2	0	0	1	1
3	1	0	2	1	2	2	1	0	2	0	0	2	1	0	1
4	1	2	0	2	1	2	0	1	0	2	0	1	2	1	0
5	2	1	2	0	1	0	2	0	1	2	1	0	1	2	0
6	2	2	1	1	0	0	0	2	2	1	1	1	0	0	2

THEOREM 7. *Let λ be a positive number, $N = 3\lambda$ and $n = 3(N-1)$. If there exists a generalized Hadamard matrix $H(\lambda, Z_3)$, then there exists an optimal $(N, 3^n)$ design D achieving the lower bound* (14) *in Theorem* 5.

The following existence results of generalized Hadamard matrices can be found implicitly in de Launey (1986) and Beth, Jungnickel and Lenz (1999).

LEMMA 1. *For any integer $k \geq 0$, there exist generalized Hadamard matrices $H(2 \cdot 3^k, Z_3)$ and $H(4 \cdot 3^k, Z_3)$.*

Combining Lemma 1 with Theorem 7, we can construct optimal $(N, 3^n)$ designs with $N = 2 \cdot 3^{k+1}$ or $4 \cdot 3^{k+1}$ and $n = 3(N-1)$ achieving the lower bound (14) in Theorem 5. Suen, Das and Midha (2013) used generalized Hadamard matrices to construct optimal fractional factorial designs. Their designs do not have minimum β-aberration whereas designs constructed from Theorem 7 do.

6. Conclusions and discussions. We introduced the concept of wordlength enumerator for general factorial designs. It is defined as an average similarity of contrasts between all possible pairs of runs. The wordlength enumerator can unify the generalized minimum aberration criterion for designs with qualitative factors and the minimum β-aberration criterion for designs with quantitative factors. We developed simple and fast methods for calculating the generalized wordlength pattern and the β-wordlength pattern with $O(N^2n^2)$ and $O(N^2n^2(s-1))$ operations, respectively, for (N, s^n) designs.

We illustrated how the wordlength enumerator can be used to rank and select designs efficiently. As an example, we searched minimum β-aberration projection designs from an OA$(36, 3^{13})$. The method is general and can be extended to deal with mixed-level designs, blocked orthogonal arrays (Lin (2014)) and split-plot orthogonal arrays (Yang and Lin (2017)) with a simple modification. Lin and Cheng (2012) examined various methods for classifying and ranking designs. Obviously, the two wordlength enumerators we considered in the paper have the same classification power as the generalized wordlength pattern and the β-wordlength pattern, respectively. Nevertheless, we can choose other y_i values to define new enumerators. For example, if we are only interested in linear and quadratic contrasts and assume that higher order contrasts are negligible, we can let $y_i = 0$ for $i = 3, \ldots, s-1$. This defines a new enumerator and a new type of wordlength pattern, which focuses on the interactions among the linear and quadratic effects. Other types of enumerators can be defined by different choices of the y_i values.

We also obtained a lower bound for three-level designs and constructed supersaturated designs that achieve the lower bound. These designs not only have generalized minimum aberration, but also have minimum β-aberration. They are particularly useful in the early stage of an investigation for screening important variables. Various analysis strategies can be used for analyzing data from such designs; see Chipman, Hamada and Wu (1997), Joseph and Delaney (2007), Yuan, Joseph and Lin (2007) and Moon, Dean and Santner (2012).

APPENDIX A: PROOFS

PROOF OF THEOREM 1. For rows a and b of $D = (d_{il})$, we have $R(d_{al}, d_{bl}) = 1 + \sum_{i=1}^{s-1} p_i(d_{al}) p_i(d_{bl}) y$, so

$$\prod_{l=1}^{n} R(d_{al}, d_{bl}) = \prod_{l=1}^{n} \left[1 + y \sum_{i=1}^{s-1} p_i(d_{al}) p_i(d_{bl}) \right],$$

which is a product of n polynomials, each with degree one. Expanding it and collecting like terms, we know that the coefficient of y^k in $\prod_{l=1}^{n} R(d_{al}, d_{bl})$ is $\sum_{(j_1,\dots,j_n)\in S_k} \prod_{l=1}^{n} p_{j_l}(d_{al}) \times p_{j_l}(d_{bl})$, where the summation is over the set $S_k = \{(j_1, \dots, j_n) : wt(j_1) + \cdots + wt(j_n) = k, 0 \le j_1, \dots, j_n \le s-1\}$. So we have

$$
\begin{aligned}
E_\alpha(D; y) &= N^{-2} \sum_{a=1}^{N} \sum_{b=1}^{N} \prod_{j=1}^{n} R(d_{aj}, d_{bj}) \\
&= N^{-2} \sum_{a=1}^{N} \sum_{b=1}^{N} \sum_{k=0}^{s-1} \left[\sum_{(j_1,\dots,j_n)\in S_k} \prod_{l=1}^{n} p_{j_l}(d_{al}) p_{j_l}(d_{bl}) \right] y^k \\
&= N^{-2} \sum_{k=0}^{s-1} \left[\sum_{(j_1,\dots,j_n)\in S_k} \sum_{a=1}^{N} \sum_{b=1}^{N} \prod_{l=1}^{n} p_{j_l}(d_{al}) p_{j_l}(d_{bl}) \right] y^k \\
&= N^{-2} \sum_{k=0}^{s-1} \sum_{(j_1,\dots,j_n)\in S_k} \left[\sum_{a=1}^{N} \prod_{l=1}^{n} p_{j_l}(d_{al}) \right]^2 y^k.
\end{aligned}
$$

Then $E_\alpha(D; y) = \sum_{k=0}^{s-1} A_k(D) y^k$ follows definition (2). $\square$

PROOF OF THEOREM 2. For rows a and b of $D = (d_{il})$, we have $R(d_{al}, d_{bl}) = 1 + \sum_{i=1}^{s-1} p_i(d_{al}) p_i(d_{bl}) y^i$, so

$$
\prod_{l=1}^{n} R(d_{al}, d_{bl}) = \prod_{l=1}^{n} [1 + p_1(d_{al}) p_1(d_{bl}) y + \cdots + p_{s-1}(d_{al}) p_{s-1}(d_{bl}) y^{s-1}],
$$

which is a product of n polynomials, each with degree $s-1$. Expanding it and collecting like terms, we know that the coefficient of y^k in $\prod_{l=1}^{n} R(d_{al}, d_{bl})$ is $\sum_{(j_1,\dots,j_n)\in S'_k} \prod_{l=1}^{n} p_{j_l}(d_{al}) \times p_{j_l}(d_{bl})$, where the summation is over the set $S'_k = \{(j_1, \dots, j_n) : j_1 + \cdots + j_n = k, 0 \le j_1, \dots, j_n \le s-1\}$. So we have

$$
\begin{aligned}
E_\beta(D; y) &= N^{-2} \sum_{a=1}^{N} \sum_{b=1}^{N} \sum_{k=0}^{n(s-1)} \left[\sum_{(j_1,\dots,j_n)\in S'_k} \prod_{l=1}^{n} p_{j_l}(d_{al}) p_{j_l}(d_{bl}) \right] y^k \\
&= N^{-2} \sum_{k=0}^{n(s-1)} \left[\sum_{(j_1,\dots,j_n)\in S'_k} \sum_{a=1}^{N} \sum_{b=1}^{N} \prod_{l=1}^{n} p_{j_l}(d_{al}) p_{j_l}(d_{bl}) \right] y^k \\
&= N^{-2} \sum_{k=0}^{n(s-1)} \sum_{(j_1,\dots,j_n)\in S'_k} \left[\sum_{a=1}^{N} \prod_{l=1}^{n} p_{j_l}(d_{al}) \right]^2 y^k.
\end{aligned}
$$

Then $E_\beta(D; y) = \sum_{k=0}^{n(s-1)} \beta_k(D) y^k$ follows definition (3). $\square$

PROOF OF THEOREM 3. Because D_1 has less β-aberration than D_2, there exists an $r \in \{1, 2, \dots, n\}$, such that $\beta_r(D_1) < \beta_r(D_2)$ and $\beta_i(D_1) = \beta_i(D_2)$ for $i = 1, \dots, r-1$. Thus

$$
E_\beta(D_1; y) - E_\beta(D_2; y) = \sum_{k=r}^{n(s-1)} \big(\beta_k(D_1) - \beta_k(D_2)\big) y^k,
$$

and

$$
\lim_{y \to 0^+} \frac{E_\beta(D_1; y) - E_\beta(D_2; y)}{y^r} = \beta_r(D_1) - \beta_r(D_2) < 0.
$$

Therefore, there exists a positive number ϵ such that $E_\beta(D_1; y) - E_\beta(D_2; y) < 0$ for all $y \in (0, \epsilon)$. $\square$

PROOF OF THEOREM 5. Based on (13), we rewrite $E(D)$ into two parts.

$$E(D) = \frac{1}{N^2}\sum_{a=1}^{N}\sigma_3^{n_3(a,a)}\sigma_4^{n_4(a,a)} + \frac{1}{N^2}\sum_{a=1}^{N}\sum_{b\neq a}\sigma_1^{n_1(a,b)}\sigma_2^{n_2(a,b)}\sigma_3^{n_3(a,b)}\sigma_4^{n_4(a,b)}.$$

As D is a balanced design, we have $\sum_{a=1}^{N} n_3(a,a) = 2Nn/3$ and $\sum_{a=1}^{N} n_4(a,a) = Nn/3$. Thus according to the inequality of arithmetic and geometric means, we have

$$\frac{1}{N}\sum_{a=1}^{N}\sigma_3^{n_3(a,a)}\sigma_4^{n_4(a,a)} \geq \left(\prod_{a=1}^{N}\sigma_3^{n_3(a,a)}\sigma_4^{n_4(a,a)}\right)^{1/N} = (\sigma_3^2\sigma_4)^{n/3}.$$

Similarly, for a balanced design, we have

$$\sum_{a=1}^{N}\sum_{b\neq a} n_1(a,b) = 4nN^2/9, \qquad \sum_{a=1}^{N}\sum_{b\neq a} n_2(a,b) = 2nN^2/9,$$

$$\sum_{a=1}^{N}\sum_{b\neq a} n_3(a,b) = 2nN(N-3)/9,$$

and

$$\sum_{a=1}^{N}\sum_{b\neq a} n_4(a,b) = nN(N-3)/9.$$

Thus according to the inequality of arithmetic and geometric means again, we have

$$\frac{1}{N(N-1)}\sum_{a=1}^{N}\sum_{b\neq a}\sigma_1^{n_1(a,b)}\sigma_2^{n_2(a,b)}\sigma_3^{n_3(a,b)}\sigma_4^{n_4(a,b)}$$

$$\geq \left(\prod_{a=1}^{N}\prod_{b\neq a}\sigma_1^{n_1(a,b)}\sigma_2^{n_2(a,b)}\sigma_3^{n_3(a,b)}\sigma_4^{n_4(a,b)}\right)^{\frac{1}{N(N-1)}} = (\sigma_1^2\sigma_2)^{\delta}(\sigma_3^2\sigma_4)^{n/3-\delta}.$$

The result is then straightforward. $\square$

PROOF OF THEOREM 6. Let $D_0 = (d_{il})$ be a saturated regular three-level design with $N = 3^k$ runs and $m = (N-1)/2$ columns. Such a design is unique up to row and column permutations. For $j = 1, 2$, let $D_j = ((d_{il} + j) \mod 3)$. Let $D = (D_0, D_1, D_2)$ be the column juxtaposed design with $n = 3(N-1)/2$ columns. We show that D satisfies the conditions in Theorem 5. First, it is easy to verify that $n_4(a,a) = n/3$ holds for any row a of D. Second, for any two distinct rows a and b, by Lemma 1 of Mukerjee and Wu (1995), there are exactly $(N-3)/6$ columns in D_0 where $d_{al} = d_{bl}$ and exactly $N/3$ columns where $d_{al} \neq d_{bl}$. Then, by the construction of D, there are exactly $(N-3)/6$ columns in D where $d_{al} = d_{bl} = 1$ so that $n_4(a,b) = (N-3)/6$, which is equal to $n/3 - \delta$, where $\delta = N/3$. Similarly, there are exactly $2 \times (N-3)/6$ columns in D where $d_{al} = d_{bl} = 0$ or 2 so that $n_3(a,b) = (N-3)/3$. Finally, for each pair of $d_{al} \neq d_{bl}$ in D_0, we will have two columns in D with $|d_{al} - d_{bl}| = 1$ and one column in D with $|d_{al} - d_{bl}| = 2$. Because there are exactly $N/3$ columns in D_0 where $d_{al} \neq d_{bl}$, so that $n_1(a,b) = 2N/3$ and $n_2(a,b) = N/3$. This shows that the quantities $n_1(a,b)$, $n_2(a,b)$, $n_3(a,b)$ and $n_4(a,b)$ for D satisfy the conditions in Theorem 5. This completes the proof. $\square$

PROOF OF THEOREM 7. Let $H = [\mathbf{0} H_0]$ be a normalized generalized Hadamard matrix $H(\lambda, Z_3)$, where $\mathbf{0}$ represents the all-zero vector. Denote $H_0 = (h_{ij})$ and $H_k = ((h_{ij} + k) \bmod 3)$ for $k = 1, 2$. Then $D = (H_0, H_1, H_2)$ is the required design. Similar to the proof of Theorem 6, by the construction of D and the properties of a generalized Hadamard matrix, we can verify that $n_4(a, a) = n/3$ for any row of D and the quantities $n_1(a, b)$, $n_2(a, b)$, $n_3(a, b)$ and $n_4(a, b)$ defined in (12) satisfy the conditions in Theorem 5. □

APPENDIX B

TABLE 6
A 36-run orthogonal array OA$(36, 3^{13})$

Row	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	1	1	1	1	1	1	1	1	1	1	1	0
3	2	2	2	2	2	2	2	2	2	2	2	2	0
4	0	0	0	0	1	1	1	1	2	2	2	2	0
5	1	1	1	1	2	2	2	2	0	0	0	0	0
6	2	2	2	2	0	0	0	0	1	1	1	1	0
7	0	0	1	2	0	1	2	2	0	1	1	2	0
8	1	1	2	0	1	2	0	0	1	2	2	0	0
9	2	2	0	1	2	0	1	1	2	0	0	1	0
10	0	0	2	1	0	2	1	2	1	0	2	1	0
11	1	1	0	2	1	0	2	0	2	1	0	2	0
12	2	2	1	0	2	1	0	1	0	2	1	0	0
13	0	1	2	0	2	1	0	2	2	1	0	1	1
14	1	2	0	1	0	2	1	0	0	2	1	2	1
15	2	0	1	2	1	0	2	1	1	0	2	0	1
16	0	1	2	1	0	0	2	1	2	2	1	0	1
17	1	2	0	2	1	1	0	2	0	0	2	1	1
18	2	0	1	0	2	2	1	0	1	1	0	2	1
19	0	1	0	2	2	2	0	1	1	0	1	2	1
20	1	2	1	0	0	0	1	2	2	1	2	0	1
21	2	0	2	1	1	1	2	0	0	2	0	1	1
22	0	1	1	2	2	0	1	0	0	2	2	1	1
23	1	2	2	0	0	1	2	1	1	0	0	2	1
24	2	0	0	1	1	2	0	2	2	1	1	0	1
25	0	2	1	0	1	2	2	0	2	0	1	1	2
26	1	0	2	1	2	0	0	1	0	1	2	2	2
27	2	1	0	2	0	1	1	2	1	2	0	0	2
28	0	2	1	1	1	0	0	2	1	2	0	2	2
29	1	0	2	2	2	1	1	0	2	0	1	0	2
30	2	1	0	0	0	2	2	1	0	1	2	1	2
31	0	2	2	2	1	2	1	1	0	1	0	0	2
32	1	0	0	0	2	0	2	2	1	2	1	1	2
33	2	1	1	1	0	1	0	0	2	0	2	2	2
34	0	2	0	1	2	1	2	0	1	1	2	0	2
35	1	0	1	2	0	2	0	1	2	2	0	1	2
36	2	1	2	0	1	0	1	2	0	0	1	2	2

Acknowledgments. The first author would like to thank Chengcheng Bo and Xiaoli Wang for their valuable discussions. The authors thank an Editor and two reviewers for their helpful comments.

Y. Tang is supported by NNSF of China with Grant 11671290.

REFERENCES

BETH, T., JUNGNICKEL, D. and LENZ, H. (1999). *Design Theory*, 2nd ed. Cambridge Univ. Press, Cambridge. MR0890103

BIRD, E. M. and STREET, D. J. (2018). Complete enumeration of all geometrically nonisomorphic three-level orthogonal arrays in 18 runs. *Australas. J. Combin.* **71** 336–350. MR3801269

CHENG, C. S. (2014). *Theory of Factorial Design: Single- and Multi-Stratum Experiments*. CRC Press/CRC, Boca Raton, FL.

CHENG, S.-W. and WU, C. F. J. (2001). Factor screening and response surface exploration. *Statist. Sinica* **11** 553–604. MR1863152

CHENG, S.-W. and YE, K. Q. (2004). Geometric isomorphism and minimum aberration for factorial designs with quantitative factors. *Ann. Statist.* **32** 2168–2185. MR2102507 https://doi.org/10.1214/009053604000000599

CHIPMAN, H., HAMADA, M. and WU, C. F. J. (1997). A Bayesian variable-selection approach for analyzing designed experiments with complex aliasing. *Technometrics* **39** 372–381.

DE LAUNEY, W. (1986). A survey of generalised Hadamard matrices and difference matrices $D(k, \lambda; G)$ with large k. *Util. Math.* **30** 5–29. MR0864807

DRAPER, N. R. and SMITH, H. (1998). *Applied Regression Analysis*. Wiley, New York. MR0211537

GEORGIOU, S. D. (2014). Supersaturated designs: A review of their construction and analysis. *J. Statist. Plann. Inference* **144** 92–109. MR3131365 https://doi.org/10.1016/j.jspi.2012.09.014

HUANG, C., LIN, D. K. J. and LIU, M.-Q. (2012). An optimality criterion for supersaturated designs with quantitative factors. *J. Statist. Plann. Inference* **142** 1780–1788. MR2903388 https://doi.org/10.1016/j.jspi.2012.02.003

JOSEPH, V. R. and DELANEY, J. D. (2007). Functionally induced priors for the analysis of experiments. *Technometrics* **49** 1–11. MR2345447 https://doi.org/10.1198/004017006000000372

LIN, D. K. J. (1993). A new class of supersaturated designs. *Technometrics* **35** 28–31.

LIN, C.-Y. (2014). Optimal blocked orthogonal arrays. *J. Statist. Plann. Inference* **145** 139–147. MR3125355 https://doi.org/10.1016/j.jspi.2013.08.016

LIN, C.-Y. and CHENG, S.-W. (2012). Isomorphism examination based on the count vector. *Statist. Sinica* **22** 1253–1272. MR2987491 https://doi.org/10.5705/ss.2010.246

LIN, C.-Y., YANG, P. and CHENG, S.-W. (2017). Minimum contamination and β-aberration criteria for screening quantitative factors. *Statist. Sinica* **27** 607–623. MR3674688

MOON, H., DEAN, A. M. and SANTNER, T. J. (2012). Two-stage sensitivity-based group screening in computer experiments. *Technometrics* **54** 376–387. MR3006387 https://doi.org/10.1080/00401706.2012.725994

MUKERJEE, R. and WU, C. F. J. (1995). On the existence of saturated and nearly saturated asymmetrical orthogonal arrays. *Ann. Statist.* **23** 2102–2115. MR1389867 https://doi.org/10.1214/aos/1034713649

MUKERJEE, R. and WU, C. F. J. (2006). *A Modern Theory of Factorial Designs*. *Springer Series in Statistics*. Springer, New York. MR2230487

PANG, F. and LIU, M.-Q. (2011). Geometric isomorphism check for symmetric factorial designs. *J. Complexity* **27** 441–448. MR2805530 https://doi.org/10.1016/j.jco.2011.04.001

SABBAGHI, A., DASGUPTA, T. and WU, C. F. J. (2014). Indicator functions and the algebra of the linear-quadratic parameterization. *Biometrika* **101** 351–363. MR3215352 https://doi.org/10.1093/biomet/ast070

SLOANE, N. J. A. (2019). Homepage of "A library of orthogonal arrays." Available at http://neilsloane.com/oadir/.

SUEN, C., DAS, A. and MIDHA, C. K. (2013). Optimal fractional factorial designs and their construction. *J. Statist. Plann. Inference* **143** 1828–1834. MR3082237 https://doi.org/10.1016/j.jspi.2013.05.004

SUN, F., LIN, D. K. J. and LIU, M.-Q. (2011). On construction of optimal mixed-level supersaturated designs. *Ann. Statist.* **39** 1310–1333. MR2816356 https://doi.org/10.1214/11-AOS877

TANG, B. and DENG, L.-Y. (1999). Minimum G_2-aberration for nonregular fractional factorial designs. *Ann. Statist.* **27** 1914–1926. MR1765622 https://doi.org/10.1214/aos/1017939244

TANG, Y. and XU, H. (2014). Permuting regular fractional factorial designs for screening quantitative factors. *Biometrika* **101** 333–350. MR3215351 https://doi.org/10.1093/biomet/ast073

WU, C.-F. J. (1993). Construction of supersaturated designs through partially aliased interactions. *Biometrika* **80** 661–669. MR1248029 https://doi.org/10.1093/biomet/80.3.661

WU, C. F. J. and HAMADA, M. S. (2009). *Experiments: Planning, Analysis, and Optimization*, 2nd ed. *Wiley Series in Probability and Statistics*. Wiley, Hoboken, NJ. MR2583259

XU, H. (2003). Minimum moment aberration for nonregular designs and supersaturated designs. *Statist. Sinica* **13** 691–708. MR1997169

XU, H., PHOA, F. K. H. and WONG, W. K. (2009). Recent developments in nonregular fractional factorial designs. *Stat. Surv.* **3** 18–46. MR2520978 https://doi.org/10.1214/08-SS040

XU, H. and WU, C. F. J. (2001). Generalized minimum aberration for asymmetrical fractional factorial designs. *Ann. Statist.* **29** 1066–1077. MR1869240 https://doi.org/10.1214/aos/1013699993

XU, H. and WU, C. F. J. (2005). Construction of optimal multi-level supersaturated designs. *Ann. Statist.* **33** 2811–2836. MR2253103 https://doi.org/10.1214/009053605000000688

YANG, P. and LIN, C.-Y. (2017). Optimal split-plot orthogonal arrays. *Aust. N. Z. J. Stat.* **59** 81–94. MR3635168 https://doi.org/10.1111/anzs.12178

YUAN, M., JOSEPH, V. R. and LIN, Y. (2007). An efficient variable selection approach for analyzing designed experiments. *Technometrics* **49** 430–439. MR2414515 https://doi.org/10.1198/004017007000000173

ZHOU, Y. and XU, H. (2015). Space-filling properties of good lattice point sets. *Biometrika* **102** 959–966. MR3431565 https://doi.org/10.1093/biomet/asv044